



A parametric study for the first-order signed integer-valued autoregressive process

Christophe Chesneau, Maher Kachour

► To cite this version:

Christophe Chesneau, Maher Kachour. A parametric study for the first-order signed integer-valued autoregressive process. *Journal of Statistical Theory and Practice*, 2012, 6 (4), pp.760–782. <hal-00643089>

HAL Id: hal-00643089

<https://hal.archives-ouvertes.fr/hal-00643089>

Submitted on 21 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A parametric study for the first-order signed integer-valued autoregressive process

Christophe Chesneau · Maher Kachour

Received:

Abstract In recent years, many attempts have been made to find accurate models for integer-valued times series. The SINAR (for Signed INteger-valued AutoRegressive) process is one of the most interesting. Indeed, the SINAR model allows negative values both for the series and its autocorrelation function. In this paper, we focus on the simplest SINAR(1) model under some parametric assumptions. Explicitly, we obtain the form the probability mass function of the innovation when the marginal distribution of the process is known. Moreover, we give an implicit form of the stationary distribution for a known innovation. Simulation experiments as well as analysis of real data sets are carried out to attest the models performance.

Keywords Integer-valued time series · INAR models · SINAR models · Rademacher(p) – \mathbb{N} class · Skellam distribution.

2000 Mathematics Subject Classification 62M10, 62M20.

1 Introduction

Integer-valued times series occur in many practical situations and several application fields. The observed data can be on \mathbb{N} , e.g. a counting process, or on \mathbb{Z} , e.g. the meteorological data. For both cases, the analysis of such a times series may present some difficulties, especially stochastic models must reflect the integer peculiarity of the observed series. For example, a standard AR(p) process does not guarantee this peculiarity. In the past decades, many attempts have been made to propose models adapted to integer-valued times series.

Christophe Chesneau
Laboratoire de Mathématiques Nicolas Oresme, Université de Caen Basse-Normandie, Campus II, Science 3, 14032 Caen, France. E-mail: chesneau@math.unicaen.fr

Maher Kachour
Laboratoire de Mathématiques Jean Leray, Université de Nantes, 2 Rue de la Houssinière, BP 92208, 44322, Nantes, France. E-mail: maher.kachour@univ-nantes.fr

Prior art on integer-valued times series on \mathbb{N} . The class of INAR (for INteger-valued AutoRegressif) models is based on the thinning operator, denoted by \circ , defined as follows. Let Z be non-negative integer-valued random variable and $a \in [0, 1]$. According to Steutel and van Harn (1979),

$$a \circ Z = \sum_{k=1}^Z \xi_k,$$

where the so-called counting sequences $(\xi_k)_{k \in \mathbb{Z}}$ are independent and identically distributed (i.i.d.) such that $\xi_1 \sim \text{Bernoulli}(a)$ i.e. $\mathbb{P}(\xi_1 = 1) = a = 1 - \mathbb{P}(\xi_1 = 0)$, and are independent of Z . The sequence $(\xi_k)_{k \in \mathbb{Z}}$ is called a counting sequences.

Based on this thinning operator, Al-zaid and Al-Osh (1987) defined an INAR(1) model mimics a real AR(1) process. In Al-zaid and Al-Osh (1990), the INAR(p) model has a correlation structure similar to the one of an ARMA(p, p-1) process. Du and Li (1991) introduced another INAR(p) process with a correlation structure identical to the one of a real AR(p) process.

Gauthier and Latour (1994) considered a more general version of the model, the GINAR(p) model, based on the generalized Steutel & van Harn operator, where the i.i.d. sequence $(\xi_k)_{k \in \mathbb{Z}}$ (the counting sequences) follows a discrete distribution with a finite mean $a \in [0, 1]$ and a finite variance μ . Hence, INAR(p) becomes a special case of GINAR(p) model. For a recent review on discrete variate time series models, we refer for example to Kedem and Fokianos (2002), McKenzie (2003), Jung and Tremayne (2006) and Weiß (2008).

However, all the previous models have the same serious limitations. Thus, because of the thinning operators, all the coefficients of the models must be positive. Therefore, the modeling of series with possible negative autocorrelations is excluded. Moreover, these models are defined on \mathbb{N} , and so they cannot fit a time series with negative observations.

Prior art on integer-valued times series on \mathbb{Z} . Integer-valued times series with negative values are very frequent in practice. But also there are other applications when values on \mathbb{Z} is the domain of interest. For example, when analyzing a non-stationary integer-valued time series with non-negative values, i.e. observations with a time trend or seasonality. In such case we need to use the differencing operator in order to achieve stationarity. Therefore, we obtain a time series on \mathbb{Z} . Another example, let (X_t) be an observation of a counting process where the values are very high, and we want to reduce the variations while keeping the integer peculiarity of the observed series. In some way, we are going to analyze the centered series (Y_t) instead of (X_t) . Note that, the empirical mean m of (X_t) is real rather than integer-valued in all but very rare cases. A mapping into the discrete support of the series is readily obtained by applying Gaussian brackets (integer part of), denoted by $[\cdot]$, or by using the rounding operator, denoted by $\langle \cdot \rangle$. Hence, one can define $Y_t = X_t - [m]$ or $Y_t = X_t - \langle m \rangle$. The observations of (Y_t) have a weak variation and they oscillate around 0 and consequently we have a time series on \mathbb{Z} .

In spite of the importance of such a modeling, it seems that few models have been proposed. A first model, we owe it to Kim and Park (2008), it is based on the signed binomial thinning operator and denoted by INARS. Then, Zhang et al. (2010) propose a generalized version of the previous model, where the counting sequences have a generalized power series as common distribution. Recently, Kachour and Truquet (2011) introduced a more general class, denoted by SINAR, based on a modified version of the generalized thinning operator, called also the signed thinning operator which is different than that introduced by Kim and Park (2008). Recall that all these processes have a autocorrelation structure identical to that of a real AR. Furthermore, consistency and the asymptotic properties of estimators proposed to estimate the parameters of these models have been well studied in the papers cited above. Note that in a different context from the thinning operation, Kachour and Yao (2009) and Kachour (2011) proposed the RINAR model, based on the rounding operator. They showed that the stationarity and ergodicity of their model is ensured under similar stationary condition as in a AR process. On the other hand, the rounding operation lead to technical difficulties: the autocorrelation structure of the RINAR(p) process is not identified, in spite of the similarity of their model with a real AR process and no asymptotic distribution is available for parameter estimates.

Contributions and relation to prior work. In the past several years, the study of the marginal distribution for INAR(1) (or GINAR(1)) type processes has become very popular. Indeed, by choosing appropriate distributions for the counting sequences and the innovation it is possible to construct a stationary first-order autoregressive integer-valued processes (INAR(1) or GINAR(1) processes) with discrete marginal distributions such as the Poisson, negative binomial, geometric, and many others (see Al-zaid and Al-Osh (1987), McKenzie (1985), McKenzie (1988) and Ristic et al. (2009a,b)). For example, consider a INAR(1) stationary process, with Bernoulli(α) as common distribution of the counting sequences, and suppose that the stationary distribution of the process is a Poisson with $\lambda/(1 - \alpha)$ as parameter. Then, we can deduce that the noise ε has also a Poisson distribution with λ as parameter. This model is called Po-INAR(1) process.

To the best of our knowledge, the question of marginal distribution for models of time series on \mathbb{Z} has often been avoided. Recently, Freeland (2010) proposed the TINAR(1) process which can be interpreted as the difference between two independent Po-INAR(1) processes. Explicitly, let $Z_t = X_t - Y_t$, where

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad Y_t = \alpha \circ Y_{t-1} + \eta_t,$$

with $\epsilon_t \sim \text{Poisson}(\lambda_1)$ and $\eta_t \sim \text{Poisson}(\lambda_2)$. Hence, the marginal distribution of Z_t is a Skellam ($\lambda_1/(1 - \alpha), \lambda_2/(1 - \alpha)$), which is the difference between two independent Poisson with respectively $\lambda_1/(1 - \alpha)$ and $\lambda_2/(1 - \alpha)$ as parameters. In particular, when $\lambda_1 = \lambda_2 = \lambda$, the marginal distribution of Z_t is symmetric around zero on \mathbb{Z} . Furthermore, note that with large values for $\mu_1 + \mu_2$, a Skellam(μ_1, μ_2) distribution can be sufficiently approximated by a gaussian distribution.

Indeed, the literature on classes of discrete distributions defined on \mathbb{Z} is very limited. The Skellam distribution, originally introduced in Skellam (1946), is one of the most interesting member. Skellam distribution has found some applications recently (see Karlis and Ntzoufras (2006, 2009) and Al-Zaid and Omair (2010)). Some other distributions defined as the difference of two discrete variables are given in Ong et al. (2008). One class of existing models are provided by the discretization of some continuous distribution such that the Gaussian distribution (see, e.g. Kemp (1997), Szablowski (2001), Roy (2003) and Inusah and Kozubowski (2006)).

In the present article we introduce a new class of distribution on \mathbb{Z} , denoted by $\text{Rademacher}(p) - \mathbb{N}$. This latter can be interpreted as the extension on \mathbb{Z} of any discrete distribution defined on \mathbb{N} , and defined as follows.

Definition 1 (*Rademacher(p) - \mathbb{N} class*) Let $p \in (0, 1)$, R and W be two independent random variables such that $R \sim \text{Rademacher}(p)$ i.e. $\mathbb{P}(R = 1) = p = 1 - \mathbb{P}(R = -1)$, and $\text{support}(W) \subseteq \mathbb{N}$. We say that a random variable X belongs to the $\text{Rademacher}(p) - \mathbb{N}$ class, if and only if

$$X \stackrel{d}{=} RW.$$

Note that, from this class and by specifying the distribution of W , we can extract many subclasses. Two of them are presented below.

Definition 2 (*R-Po(p, λ) and SPo(λ) classes*) Let $p \in (0, 1)$ and $\lambda > 0$. We define the Rademacher-Poisson class, and we denote it by $\text{R-Po}(p, \lambda)$, all the random variables X belonging to the $\text{Rademacher}(p) - \mathbb{N}$ class where $W \sim \text{Poisson}(\lambda)$. The $\text{R-Po}(1/2, \lambda)$ is called the symmetric-Poisson class and denoted by $\text{SPo}(\lambda)$.

The main purpose of this paper is to introduce and develop a new family of SINAR(1) process, by considering a parametric assumption on the common distribution of the counting sequence of the model. The new model will be denoted by P-SINAR(1). Thus, using the characteristic function, we give a full form of innovation distribution when the marginal distribution of the new process belongs to the $\text{Rademacher}(p) - \mathbb{N}$ class. Then, we extend the previous result by giving an implicit form of the innovation distribution when the marginal distribution of P-SINAR(1) is known (not necessarily a member of the $\text{Rademacher}(p) - \mathbb{N}$ class, e.g. Skellam). Furthermore, we obtain an implicit form of the marginal distribution of P-SINAR(1) process when the innovation distribution is known (e.g. symmetric Poisson, Skellam). Note that these results are very useful for the for prediction purposes and the goodness of fit of the P-SINAR(1) model. Namely, identifying the noise distribution allows to calculate the transition probability of the process, which is simply the one step ahead predictive distribution. On the other hand, knowing the marginal distribution of P-SINAR(1) allows to verify if given real data can be adjusted by the model.

Paper organization. The rest of this paper is organized as follows. In section 2, the first-order Parametric Signed INteger-valued AutoRegressive

(P-SINAR(1)) model is presented together with some of its properties. Our main results are set in Section 3. A one-step ahead forecasting approach is described in Section 4. Simulation experiments as well as analysis of real data are discussed in Section 5 and 6. The proofs of all forthcoming results are postponed to Section 7.

2 The P-SINAR(1) process

The so-called signed thinning operator, originally proposed by Latour and Truquet (2008), seems to be the natural extension of the classical Steutel & Van Harn operator to \mathbb{Z} -valued random variables and is defined as follows.

Definition 3 (*Signed thinning operator*) Let $(Y_i)_{i \in \mathbb{Z}}$ be a sequence of i.i.d. integer-valued random variables with F as common distribution, independent of an integer-valued random variable X . The signed thinning operator, denoted by $F \circ$, is defined by

$$F \circ X = \begin{cases} \text{sign}(X) \sum_{i=1}^{|X|} Y_i, & \text{if } X \neq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where for an integer x , $\text{sign}(x) = 1$ if $x > 0$ and -1 if $x < 0$. The sequence $(Y_i)_{i \in \mathbb{Z}}$ is referred to as a counting sequence.

Based on the above operator, Kachour and Truquet (2011) introduced the SINAR (for Signed INteger-valued AutoRegressive) process, which can be considered as a natural extension on \mathbb{Z} of the well-known INAR (or GINAR) process. In this paper, we focus on the SINAR(1) model defined as follows.

Definition 4 (*SINAR(1) process*) A sequence $(X_t)_{t \in \mathbb{Z}}$ is said to be a SINAR(1) process if it admits the following representation

$$X_t = F_1 \circ X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (2)$$

where $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a sequence of i.i.d. integer-valued random variables, with mean μ_ε and variance σ_ε^2 . It is understood that the counting sequence $Y_t^{(1)}$ associated to the operator $F_1 \circ$ is independent of ε_t . Moreover, ε_t and the sigma-algebra $\mathcal{G}_{t-1} = \sigma(X_{t-1}, \dots)$, $t \in \mathbb{Z}$, are supposed to be mutually independent.

Next, we summarize some theoretical results concerning stationarity ergodicity and autocorrelation structure of the SINAR(1) process (for details, we refer to Kachour and Truquet (2011)). First, we recall the following assumptions.

(A1) F_1 charges zero, i.e. $F_1\{0\} = \mathbb{P}(Y_1 = 0) > 0$.

(A1') There exist an even integer $a \in \mathbb{Z}$ and an odd integer $b \in \mathbb{Z}$ such that

$$\mathbb{P}(\varepsilon_1 = a) > 0, \quad \mathbb{P}(\varepsilon_1 = b) > 0.$$

(A2) We have $|\mathbb{E}(Y_1)| < 1$.

Suppose that **(A1)** (or **(A1')**) and **(A2)** hold. Then the equation (2) has a unique stationary solution X satisfying $\mathbb{E}(X^2) < \infty$. Finally, note that the process X is also ergodic.

Remark 1 A stationary SINAR(1) process has the same autocorrelation structure than a classical AR(1) process

$$Y_t = \alpha_1 Y_{t-1} + \eta_t, \quad t \in \mathbb{Z},$$

where α_1 denotes the mean of the distribution F_1 (which is not necessarily positive) and η_t denotes a white noise. In other words, we have

$$\text{cor}(X_t, X_{t+k}) = \alpha_1^{|k|}.$$

Now, we introduce a new family of SINAR(1) process, by specifying F_1 the common distribution of the counting sequence of the process. The new model will be denoted by P-SINAR(1) process (for Parametric Signed INteger-valued AutoRegressive).

Definition 5 (*P-SINAR(1) process*) A sequence $(X_t)_{t \in \mathbb{Z}}$ is said to be a P-SINAR(1) process if it admits the same representation of a SINAR(1) process, i.e.

$$X_t = F_1 \circ X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (3)$$

with F_1 the common distribution of the i.i.d. counting sequence $(Y_i)_{i \in \mathbb{Z}}$ defined as follows:

$$\mathbb{P}(Y_1 = -1) = (1 - \alpha)^2, \quad \mathbb{P}(Y_1 = 0) = 2\alpha(1 - \alpha), \quad \mathbb{P}(Y_1 = 1) = \alpha^2, \quad (4)$$

where $\alpha \in (0, 1)$.

Remark 2 We have $Y_1 \stackrel{d}{=} Z_1 - 1$, where $Z_1 \sim \text{Binomial}(2, \alpha)$. Moreover, for all $k \in \mathbb{N}^*$, we have

$$\mathbb{P}\left(\sum_{i=1}^k Y_i = l\right) = \mathbb{P}(Z_k = k + l), \quad l \in \{-k, \dots, k\}, \quad (5)$$

where $Z_k \sim \text{Binomial}(2k, \alpha)$.

In the next theorem, we derive some properties of the P-SINAR(1) process.

Theorem 1 Consider a P-SINAR(1) model, as defined in (3). Then,

1. The process has a unique stationary solution, denoted by X , and it is also ergodic.

2. (X_t) forms an homogeneous Markov chain with state space $E = \mathbb{Z}$ and its transition probability function is given by

$$\mathbb{P}(X_t = a \mid X_{t-1} = 0) = \mathbb{P}(\varepsilon = a), \quad a \in \mathbb{Z}, \quad (6)$$

and, for all $a \in \mathbb{Z}$ and $b \in \mathbb{Z}^*$,

$$\mathbb{P}(X_t = a \mid X_{t-1} = b) = \sum_{l=-|b|}^{|b|} \mathbb{P}(\varepsilon = a - l) \mathbb{P}(Z_{|b|} = |b| + \text{sign}(b)l), \quad (7)$$

where $Z_{|b|} \sim \text{Binomial}(2|b|, \alpha)$.

Moreover, under second order stationarity conditions, we obtain

$$\mathbb{E}(X_t) = \frac{\mu_\varepsilon}{1 - \alpha_1}, \quad \mathbb{V}(X_t) = \frac{\beta_1 \mathbb{E}(|X_t|) + \sigma_\varepsilon^2}{1 - \alpha_1^2}, \quad (8)$$

where

$$\alpha_1 = \mathbb{E}(Y_1) = 2\alpha - 1, \quad \beta_1 = \mathbb{V}(Y_1) = 2\alpha(1 - \alpha) = \mathbb{P}(Y_1 = 0).$$

3 Results

3.1 Main results

First, some recalls and notations are necessary. For any random variable U , let ϕ_U be the characteristic function of U and G_U its probability generating function i.e.

$$\phi_U(t) = \mathbb{E}(e^{itU}), \quad t \in \mathbb{R}, \quad G_U(z) = \mathbb{E}(z^U), \quad |z| < 1,$$

Remark 3 When U has a distribution on \mathbb{Z} , only the characteristic function is useful to characterize the distribution: it exists for all probability distributions and it determines uniquely the law of a random variable in the sense that $\phi_U = \phi_V$ is equivalent to " U and V have the same distribution".

In Theorem 2 below, we give the form of the characteristic function of the noise $\varepsilon \stackrel{d}{=} \varepsilon_1$ of a P-SINAR(1) process with a stationary distribution belonging to the Rademacher(p) – \mathbb{N} class.

Theorem 2 Consider a P-SINAR(1) process as defined in (3). Suppose that its marginal stationary distribution, denoted by X , belongs to the Rademacher(p) – \mathbb{N} class, i.e. $X \stackrel{d}{=} RW$, where R and W are two independent random variables such that $R \sim \text{Rademacher}(p)$, $p \in (0, 1)$, and $\text{support}(W) \subseteq \mathbb{N}$. Then, the characteristic function of the noise ε is

$$\phi_\varepsilon(t) = \frac{p\phi_W(t) + (1-p)\phi_W(-t)}{pG_W(\phi_Y(t)) + (1-p)G_W(\phi_Y(-t))}, \quad t \in \mathbb{R},$$

where

$$\phi_Y(s) = 2\alpha(1 - \alpha) + \alpha^2 e^{is} + (1 - \alpha)^2 e^{-is}, \quad s \in \mathbb{R}.$$

Moreover, we have

$$\mathbb{P}(\varepsilon = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{\varepsilon}(t) e^{-ikt} dt, \quad k \in \mathbb{Z}. \quad (9)$$

Remark 4 Unfortunately, even if W is known (e.g. $W \sim \text{Poisson}(\lambda)$), we can not confirm that ε has a known distribution of \mathbb{Z} . Indeed, this is not a real surprise, since the known classes of distributions on \mathbb{Z} are very limited and often such classes are rather built.

In the next theorem, we give an implicit form of the distribution of the noise ε of the P-SINAR(1) process (depending on the marginal stationary distribution of (3)).

Theorem 3 Consider a P-SINAR(1) process as defined in (3). Suppose that X its marginal distribution is known (not necessarily a member of the Rademacher(p) – \mathbb{N} class). Then, we have

$$\mathbb{P}(\varepsilon = k) = \frac{1}{1 - \mathbb{P}(X = 0)} (Q_1(k) + Q_2(k)), \quad k \in \mathbb{Z}, \quad (10)$$

where

$$Q_1(k) = \sum_{m=1}^{\infty} \mathbb{P}(X = m) \sum_{j=k-m}^{k+m} \sum_{l=-m}^m \mathbb{P}(\varepsilon = j-l) \mathbb{P}(Z_m = m+j-k) \mathbb{P}(Z_m = m+l) \quad (11)$$

and

$$Q_2(k) = \sum_{m=1}^{\infty} \mathbb{P}(X = -m) \sum_{j=k-m}^{k+m} \sum_{l=-m}^m \mathbb{P}(\varepsilon = j-l) \mathbb{P}(Z_m = m-j+k) \mathbb{P}(Z_m = m-l), \quad (12)$$

with $Z_m \sim \text{Binomial}(2m, \alpha)$.

Remark 5 Theorem 3 seems more general than Theorem 2, since we do not specifying the distribution of X . On the other hand, the implicit form of the distribution of the noise ε , given by the equations (10), (11) and (12), is difficult to exploit even numerically. However, supposing that the support of X is finite (which means that the stationary process (3) is an homogeneous Markov chain with a finite state space $E = \text{support}(X)$), the calculation becomes feasible. For example, let $\text{support}(X) = \{-M, \dots, M\}$, with $M \in \mathbb{N}^*$. From the definition of signed thinning operator (1) and under Assumption [H], we have

$$\text{support}(\varepsilon) = \{-2M, \dots, 2M\}.$$

Therefore, (10), (11) and (12) lead to a general system of $2M$ linear equations with $2M$ unknowns. Moreover, since $\sum_{k=-2M}^{2M} \mathbb{P}(\varepsilon = k) = 1$, this system has a unique solution (even if M is very large) and then the distribution of ε can be calculated.

In the sequel, we give an implicit form of the marginal stationary distribution X of the P-SINAR(1) process (depending on the distribution of the noise ε).

Theorem 4 *Consider a P-SINAR(1) process as defined in (3). Suppose that the distribution of the noise ε is known. Then, the marginal stationary distribution, denoted by X , of (3) is given as follows.*

For all $k \in \mathbb{Z}$, we have

$$\mathbb{P}(X = k) = \mathbb{P}(F_1 \circ X = 0)\mathbb{P}(\varepsilon = k) + \sum_{j \in \mathbb{Z}^*} \mathbb{P}(F_1 \circ X = j)\mathbb{P}(\varepsilon = k - j), \quad (13)$$

where, for all $l \in \mathbb{Z}$,

$$\begin{aligned} \mathbb{P}(F_1 \circ X = l) = & \sum_{k=1}^{\infty} (\mathbb{P}(Z_k = k + l)\mathbb{P}(X = k) + \mathbb{P}(Z_k = k - l)\mathbb{P}(X = -k)) \mathbf{1}_{\{l \in \{-k, \dots, k\}\}} \\ & + \mathbb{P}(X = 0)\mathbf{1}_{\{l=0\}}, \end{aligned}$$

with $Z_k \sim \text{Binomial}(2k, \alpha)$.

Remark 6 Note that the implicit form of the distribution of X given by the Theorem 4 is also difficult to handle. However, additional assumptions may reduce the calculation, e.g. a noise with a finite support, without making it really feasible.

3.2 Estimation

In this section, we describe estimation for the parameters of the P-SINAR(1) process under some parametric additional assumptions. Here, we recall that, from (4), we have

$$\mathbb{E}(Y_1) = \alpha_1 = 2\alpha - 1, \quad \mathbb{V}(Y_1) = \beta_1 = 2\alpha(1 - \alpha). \quad (14)$$

Assumption on X the marginal distribution of the process. Suppose that X follows a symmetric-Poisson distribution $\text{SPo}(\lambda)$, i.e. $X \stackrel{d}{=} RW$ where $R \sim \text{Rademacher}(1/2)$ and $W \sim \text{Poisson}(\lambda)$. Since $\mathbb{E}(X) = 0$ and from (8), we have that μ_ε the mean of the noise ε equals zero. Let X_0, X_1, \dots, X_n be observations from the P-SINAR(1) process with the previous assumption. Thus, the conditional least squares (CLS) estimator of α_1 (mean of F_1) is defined by

$$\hat{\alpha}_n = \frac{\sum_{t=1}^n X_{t-1} X_t}{\sum_{t=1}^n X_{t-1}^2}. \quad (15)$$

Moreover, from (14), one can deduce that the CLS estimator of α (the actual parameter of the distribution F_1), is given by

$$\hat{\hat{\alpha}}_n = \frac{1 + \hat{\alpha}_n}{2}.$$

Theorem 5 Consider a P -SINAR(1) process as defined in (3). Assume that $X \sim \text{SPo}(\lambda)$. Then, $\hat{\alpha}_n$ is strongly consistent, i.e.

$$\lim_{n \rightarrow \infty} \hat{\alpha}_n = \alpha_1, \quad a.s.$$

and

$$\sqrt{n}(\hat{\alpha}_n - \alpha_1) \xrightarrow{d} \mathcal{N}(0, \Psi), \quad (16)$$

where

$$\Psi = 2\alpha(1 - \alpha) \frac{2\lambda^3 + 4\lambda^2 + 4\lambda + 1}{\lambda(1 + \lambda)^2}.$$

On the other hand, since $\mathbb{P}(X = 0) = e^{-\lambda}$, the parameter λ of the marginal stationary distribution can be estimated by

$$\hat{\lambda}_n = \log \left(\frac{n}{\sum_{i=1}^n \mathbf{1}_{\{X_i=0\}}} \right). \quad (17)$$

Note that, by using the ergodicity of the process (3), one can prove that $\hat{\lambda}_n$ is strongly consistent, and has a gaussian asymptotic behavior, i.e. $\sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow{d} \mathcal{N}(0, \Phi)$ where $\Phi > 0$.

Remark 7 Let X be a random variable following a R -Po(p, λ) distribution, i.e. $X \stackrel{d}{=} RW$ where $R \sim \text{Rademacher}(p)$ and $W \sim \text{Poisson}(\lambda)$. Then, the parameter λ can be again estimated by (17). Moreover, from (8), the parameter p can be estimated by

$$\hat{p}_n = \frac{1}{2} \left(\frac{\hat{\mu}_n}{(1 - \hat{\alpha}_n)\hat{\lambda}_n} + 1 \right),$$

where $\hat{\mu}_n$ denotes the CLS estimator of μ_ε (the noise mean), which has also a gaussian asymptotic behavior and defined by $\hat{\mu}_n = \bar{X}_{n,0} - \hat{\alpha}_n \bar{X}_{n,1}$, with $\bar{X}_{n,j} = (1/n) \sum_{t=1}^n X_{t-j}$, $j \in \{0, 1\}$.

Remark 8 Consider now that X follows a symmetric Skellam distribution, i.e. $X = Z_1 - Z_2$, where Z_1 and Z_2 are i.i.d. with $Z_1 \sim \text{Poisson}(\lambda)$. The conditional least squares estimator $\hat{\alpha}_n$ still has gaussian asymptotic behavior. In this case, the asymptotic variance is more complicated to determine analytically, since it involves calculating $\mathbb{E}(|X|)$ and $\mathbb{E}(|X|^3)$, which can be deduced as in Katti (1960) but there are not in a helpful form. On the other hand, since the marginal variance is 2λ , the parameter λ can be estimated by the sample variance as follows:

$$\hat{\lambda}_n = \frac{1}{2n} \sum_{t=1}^n X_t^2. \quad (18)$$

Assumption on the noise distribution. Suppose that ε in (3) follows a symmetric-Poisson distribution $\text{SPo}(\lambda)$. Therefore, we have $\mu_\varepsilon = 0$ and $\sigma_\varepsilon^2 = \lambda(1 + \lambda)$. Hence, the conditional least squares will yield only an estimate for α_1 and not for λ . Let X_0, X_1, \dots, X_n be observations from the P-SINAR(1) process with the previous assumption. Here, we will use the moment estimation which is simple and straightforward based on (8) and the fact that $\lambda > 0$. Let $\hat{\rho}_n(1)$ denotes the first order sample autocorrelation function, s^2 the empirical variance, and

$$|\bar{x}| = \frac{1}{n} \sum_{t=1}^n |X_t|.$$

Then, it follows that

$$\hat{\alpha}_1 = \hat{\rho}_n(1), \quad \hat{\alpha} = \frac{1 + \hat{\alpha}_1}{2}, \quad (19)$$

and

$$\hat{\lambda} = \frac{\sqrt{1 + 4\kappa_n} - 1}{2}, \quad (20)$$

where

$$\kappa_n = 2\hat{\alpha}(1 - \hat{\alpha})(2s^2 - |\bar{x}|). \quad (21)$$

Remark 9 Suppose now that ε follows a symmetric Skellam distribution $\text{SKellam}(\lambda)$. Thus, we have $\mu_\varepsilon = 0$ and $\sigma_\varepsilon^2 = 2\lambda$. Therefore, the moment estimates of the parameters are defined as follows:

$$\hat{\alpha}_1 = \hat{\rho}_n(1),$$

and

$$\hat{\lambda} = \frac{\kappa_n}{2}, \quad (22)$$

where κ_n is given by (21).

Remark 10 Since the distribution of ε is known, others estimation methods can be considered. Namely, the conditional maximum likelihood (CML) estimators of the parameters are obtained by maximizing the function

$$L_\theta(x) = \prod_{t=2}^n \mathbb{P}(X_t = x_t \mid X_{t-1} = x_{t-1}),$$

where $\theta = (\alpha_1, \lambda)$, $x = (x_1, \dots, x_n) \in \mathbb{Z}^n$, and the conditional distribution $X_t \mid X_{t-1}$ is given by (6) and (7) of Theorem 1. Note that, the CML estimation is more complicated (but tractable) and their results are slightly better than those obtained by the moments estimation method described above.

4 Forecasting

In this section, we focus on one step ahead predictions for the P-SINAR(1) process, as defined in (3). The one-step least squares ahead forecast \tilde{X}_{T+1} of X_{T+1} equals

$$\tilde{X}_{T+1} = \mathbb{E}(X_{T+1} \mid X_T) = \alpha_1 X_T + \mu_\varepsilon,$$

where $\alpha_1 = 2\alpha - 1$ denotes the mean of F_1 and μ_ε that of the noise. In general \tilde{X}_{T+1} is real-valued. In order to get integer values we can use the median from the transition probability function of the process, defined in (6) and (7), together with its 95% confidence limits, i.e.

$$\sum_{j=M}^N \mathbb{P}(X_{T+1} = j \mid X_T = i) \geq 0.95$$

such as $N - M$ to be as small as possible. Note that, the conditional distribution $X_{T+1} \mid X_T$ depends on the noise distribution.

Remark 11 Another mapping into the discrete support of the series can be obtained by rounding to the nearest integer. Thus, we find

$$\tilde{X}_{T+1} = \langle \alpha_1 X_T + \mu_\varepsilon \rangle, \quad (23)$$

where $\langle \cdot \rangle$ denotes the rounding operator. Note that this method is not very efficient. For instance, suppose that we consider two P-SINAR(1) models with different centered noises (e.g. $\varepsilon_t \sim \text{SPo}(\lambda)$ and $\eta_t \sim \text{Skellam}(\lambda)$) to fit a given real data. According to (23), in spite of their difference, both models have the same prediction performances.

5 Numerical illustration

In the sequel, we consider a P-SINAR(1) process, as defined in (3), with a fixed $\alpha = 0.8$. It follows that,

$$\mathbb{P}(Y_1 = -1) = 0.04, \quad \mathbb{P}(Y_1 = 0) = 0.32, \quad \mathbb{P}(Y_1 = 1) = 0.64.$$

First numerical study: on the distribution of ε . We suppose that $\text{support}(X) = \{-2, -1, 0, 1, 2\}$, and its distribution is given by

$$\begin{aligned} \mathbb{P}(X = 0) &= 0.7, & \mathbb{P}(X = 1) &= \mathbb{P}(X = -1) = 0.14, \\ \mathbb{P}(X = 2) &= \mathbb{P}(X = -2) = 0.01. \end{aligned} \quad (24)$$

According to Remark 5, we have $\text{support}(\varepsilon) = \{-4, \dots, 4\}$ and its distribution is identified by solving the following system of equations: for all $k \in \{-4, \dots, 4\}$,

$$\mathbb{P}(\varepsilon = k) = \sum_{m=1}^2 \mathbb{P}(X = m) \sum_{j=k-m}^{k+m} \left(\sum_{l=-m}^m \mathbb{P}(\varepsilon = j - l) H(j, l, m, k) \right),$$

where

$$\begin{aligned} H(j, l, m, k) &= \mathbb{P}(Z_m = m + j - k) \mathbb{P}(Z_m = m + l) \\ &\quad + \mathbb{P}(Z_m = m - j + k) \mathbb{P}(Z_m = m - l) \\ &= \binom{2m}{m+l} \binom{2m}{m+j-k} G(j, l, m, k), \end{aligned}$$

with

$$G(j, l, m, k) = 0.8^{2m+(j-k+l)} 0.2^{2m-(j-k+l)} + 0.8^{2m-(j-k+l)} 0.2^{2m+(j-k+l)}.$$

In Figure 1, we depict the probability mass function of ε distribution for this case.

Remark 12 The histogram of the real data, studied in Section 6 below, is consistent with the specific distribution of X defined in (24).

Second numerical study: on the distribution of X . Now, we suppose that the distribution of ε is known. Hence, by Theorem 4, the marginal stationary distribution X of the process (3) has an implicit form given by (13). Next, we consider two types of distribution for ε :

- First, we suppose that ε follows a symmetric Skellam distribution distribution with parameter λ , i.e. $X = Z_1 - Z_2$, where Z_1 and Z_2 are i.i.d. with $Z_1 \sim \text{Poisson}(\lambda)$. Recall that with large values for λ , a symmetric Skellam(λ) distribution can be sufficiently approximated by a centered gaussian distribution.

In Figure 2, we depict the probability mass function of the marginal stationary distribution of the P-SINAR(1) for $\lambda \in \{1, 2, 3, 7, 10, 30\}$. Hence, we remark that with large values for λ , X can be also sufficiently approximated by a centered gaussian distribution.

- Now, we suppose that ε follows a symmetric-Poisson distribution. Recall that, Symmetric-Poisson is a subclass of Rademacher(p) – \mathbb{N} class (see Definition 2). Moreover, for large values of λ , a R-Po(p, λ) distribution can be sufficiently approximated by two gaussian distributions, one on \mathbb{N} and the other on $\mathbb{Z}^- = \mathbb{Z} \setminus \mathbb{N}$ (details are given in Annex).

In Figure 3, we depict the probability mass function of the marginal stationary distribution of the SINAR(1) for $\lambda \in \{1, 2, 3, 7, 10, 30\}$. Thus, we find that X has a gaussian approximation with large values of λ .

6 Application to real data

The models introduced in this paper are applied to 597 counts of the price change for the Australian firm Broken Hill Proprietary (BHP) Limited, measured in cents, at each transaction time during the day of October 1, 2001. Note

that, the observations vary from -2 to 2 , the sample mean equals -0.00335 and the variance 0.41274 .

In Figure 4, we have the time series plot, histogram, sample autocorrelation function and sample partial autocorrelation function. Thus, the correlogram shows that the first sample autocorrelation, is more significant than the others and there exists a cutt-off after lag 1 in the partial autocorrelations. That justify an AR(1) process fitting. Recall that a SINAR(1) process is defined on \mathbb{Z} , has a correlation structure similar to a real AR(1), and allows negative autocorrelation. Thus, for the previous reasons, one can consider a SINAR(1) to analyse the data. On the other hand, according to the nature of the observations, assuming that the distribution F_1 is defined by (4) does not seem abusive. Therefore, in the sequel we will use a P-SINAR(1) to fit the data.

Furthermore, since the observations are centered, it is natural to consider a mean zero noise. For our first attempt to fit the data, we propose a P-SINAR(1) model such that the noise follows a symmetric Poisson distribution, i.e. $\varepsilon \sim \text{SPo}(\lambda_1)$. First, we estimated the model parameters using (19), (20) and (21). It follows that

$$\hat{\alpha}_1 = -0.2623 \quad \left(\hat{\alpha} = \frac{1 + \hat{\alpha}_1}{2} = 0.3688 \right),$$

and

$$\hat{\lambda}_1 = 0.1859.$$

Recall that, according to Theorem 4, the marginal stationary distribution of the process has an implicit form given by (13), which is difficult to handle. Thus, in order to assess the goodness of fit we followed a computational approach. Explicitly, using the above estimated parameter values, we simulated 1000 series of length 597 and we kept the expected relative frequencies for each series. The reported frequencies are the mean over the 1000 series.

The Table 1 below contains the observed and expected relative frequencies based on the proposed model. One can see that the model fits very well the data. In Figure 5, we depict the one ahead predictive distribution $X_{t+1} \mid X_t = i$ for values of $i = 1, -1, 2, -2$.

Now, we will calculate through simulation the standard errors of the moment estimation method considered to estimate the parameters of the above proposed model. Again, by using the estimated parameters, i.e. $\hat{\alpha} = 0.3688$ and $\hat{\lambda} = 0.1859$, we simulate a serie of length 597, and we find the moment estimates. After $K = 1000$ independent replications, we calculate the average and standard deviation of $\{\hat{\alpha}_i, i = 1, \dots, K\}$ and $\{\hat{\lambda}_i, i = 1, \dots, K\}$ the sequences of the obtained estimates. This resulted in $\hat{\alpha} = 0.3677$ and $\hat{\lambda}_1 = 0.1867$ with standard errors of 0.0296 and 0.0257, respectively.

Note that, one can propose a P-SINAR(1) model with a symmetric Skellam distribution for the noise, i.e. $\varepsilon \sim \text{Skellam}(\lambda_2)$. Then, we estimated the model parameters using (19), (21) and (22). Hence, we obtain

$$\hat{\alpha}_1 = -0.2623, \text{ and } \hat{\lambda}_2 = 0.1103.$$

Again, in order to access the goodness of fit, we proceed in the same manner as the previous case.

Table 2 contains the observed and expected relative frequencies based on the new proposed model. One can see that the new model fits also very well the data. We depict the one ahead predictive distribution $X_{t+1} | X_t = i$ for values of $i = 1, -1, 2, -2$, in Figure 6.

Remark 13 The estimated values of λ_1 and λ_2 are quite close and weak (both < 0.2). Furthermore, remark that Skellam and Symmetric Poisson distributions have similar behavior with weak parameter values. This explain that the one ahead predictive distributions for both above models (displayed in Figure 5 and 6) are also very close. Finally, we note that the one step ahead prediction (for both models) will often leads to zero.

Remark 14 The histogram of the data (see Figure 4) is consistent with a symmetric Skellam distribution. Therefore, one can also propose a P-SINAR(1) process, where its marginal distribution $X \sim \text{Skellam}(\lambda_3)$, to fit the data. The parameters of the model, i.e. α_1 (or α) and λ_3 , can be estimated by (15) and (18). According to Theorem 3, the noise distribution of the process has an implicit form given by (10) (11) and (12), which is difficult to exploit analytically. However, using the data, one can simulated the noise distribution.

Values	Observed	Expected
≤ -2	0.0134	0.0217
-1	0.1507	0.1557
0	0.67839	0.6700
1	0.1407	0.1373
≥ 2	0.0167	0.01507

Table 1 Observed versus expected probabilities, when $\varepsilon \sim \text{SPo}$

Values	Observed	Expected
≤ -2	0.0134	0.0167
-1	0.1507	0.149
0	0.67839	0.6448
1	0.1407	0.1675
≥ 2	0.0167	0.0216

Table 2 Observed versus expected probabilities, when $\varepsilon \sim \text{Skellam}$

7 Proofs

Proof of Theorem 1

Stationarity of P-SINAR(1). Recall that, we have

$$F_1\{0\} = \mathbb{P}(Y_1 = 0) = 2\alpha(1 - \alpha) > 0$$

and

$$|\mathbb{E}(F_1)| = |2\alpha - 1| < 1.$$

Thus, assumptions **(A1)** and **(A2)** of Theorem 1 of Kachour and Truquet (2011) are verified (they have been recalled in Section 2), and then the P-SINAR(1) process (3) has a unique stationary solution, denoted by X , and it is also ergodic.

Markovian properties of the process. Since the noise $(\varepsilon_t)_{t \in \mathbb{Z}}$ in (3) is an i.i.d. sequence of integer-valued random variables, then the P-SINAR(1) process, denoted by $(X_t)_{t \in \mathbb{Z}}$, is Markovian. Thus, from the definition of the signed thinning operator, for all $a \in \mathbb{Z}$, we have (with $\varepsilon \stackrel{d}{=} \varepsilon_t$)

$$\Gamma(a, 0) = \mathbb{P}(X_t = a \mid X_{t-1} = 0) = \mathbb{P}(\varepsilon = a). \quad (25)$$

Furthermore, for any $a \in \mathbb{Z}$ and $b \in \mathbb{Z}^*$, we obtain

$$\begin{aligned} \Gamma(a, b) &= \mathbb{P}(X_t = a \mid X_{t-1} = b) \\ &= \mathbb{P}\left(\text{sign}(b) \sum_{k=1}^{|b|} Y_k + \varepsilon_t = a\right) \\ &= \sum_{l \in \mathbb{Z}} \mathbb{P}(\varepsilon = a - l) \mathbb{P}\left(\sum_{k=1}^{|b|} Y_k = \text{sign}(b)l\right). \end{aligned}$$

Recall that, from (5), for any $k \in \mathbb{N}^*$, we have

$$\mathbb{P}\left(\sum_{i=1}^k Y_i = l\right) = \mathbb{P}(Z_k = k + l), \quad l \in \{-k, \dots, k\},$$

where $Z_k \sim \text{Binomial}(2k, \alpha)$.

Therefore, we find

$$\Gamma(a, b) = \sum_{l=-|b|}^{|b|} \mathbb{P}(\varepsilon = a - l) \mathbb{P}(Z_{|b|} = |b| + \text{sign}(b)l). \quad (26)$$

Theorem 1 is proved. ■

Proof of Theorem 2

Recall that for any random variable U , we denote by ϕ_U its characteristic function and G_U its probability generating function.

By the stochastic representation of the model (3), we have (with $X \stackrel{d}{=} X_t$)

$$\phi_\varepsilon(t) = \frac{\phi_X(t)}{\phi_{F_1 \circ X}(t)}. \quad (27)$$

Calculation of $\phi_X(t)$. We have

$$\phi_X(t) = \mathbb{P}(X = 0) + \mathbb{E}(e^{itX} \mathbf{1}_{\{X > 0\}}) + \mathbb{E}(e^{itX} \mathbf{1}_{\{X < 0\}}).$$

Observe that

$$\begin{aligned} \mathbb{E}(e^{itX} \mathbf{1}_{\{X > 0\}}) &= \sum_{k=1}^{\infty} e^{itk} \mathbb{P}(X = k) = p \sum_{k=1}^{\infty} e^{itk} \mathbb{P}(W = k) \\ &= p \left(\sum_{k=0}^{\infty} e^{itk} \mathbb{P}(W = k) - \mathbb{P}(W = 0) \right) \\ &= p\phi_W(t) - p\mathbb{P}(X = 0) \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}(e^{itX} \mathbf{1}_{\{X < 0\}}) &= \sum_{k=-\infty}^{-1} e^{itk} \mathbb{P}(X = k) = (1-p) \sum_{k=-\infty}^{-1} e^{itk} \mathbb{P}(W = -k) \\ &= (1-p) \left(\sum_{k=0}^{\infty} e^{-itk} \mathbb{P}(W = k) - \mathbb{P}(W = 0) \right) \\ &= (1-p)\phi_W(-t) - (1-p)\mathbb{P}(X = 0). \end{aligned}$$

Therefore

$$\phi_X(t) = p\phi_W(t) + (1-p)\phi_W(-t). \quad (28)$$

Calculation of $\phi_{F_1 \circ X}(t)$. By the definition of the signed thinning operator $F_1 \circ$, we have

$$\phi_{F_1 \circ X}(t) = \mathbb{P}(X = 0) + \mathbb{E} \left(e^{it \sum_{i=1}^X Y_i} \mathbf{1}_{\{X > 0\}} \right) + \mathbb{E} \left(e^{-it \sum_{i=1}^{-X} Y_i} \mathbf{1}_{\{X < 0\}} \right).$$

It follows from the i.i.d. nature of $(Y_i)_{i \in \mathbb{Z}}$, the independence of $(Y_i)_{i \in \mathbb{Z}}$ and X , and $X \stackrel{d}{=} RW$ that

$$\begin{aligned} \mathbb{E} \left(e^{it \sum_{i=1}^X Y_i} \mathbf{1}_{\{X > 0\}} \right) &= \mathbb{E} \left(\mathbf{1}_{\{X > 0\}} \mathbb{E} \left(e^{it \sum_{i=1}^X Y_i} \mid X \right) \right) \\ &= \sum_{k=1}^{\infty} (\phi_Y(t))^k \mathbb{P}(X = k) \\ &= p \sum_{k=1}^{\infty} (\phi_Y(t))^k \mathbb{P}(W = k) \\ &= p \left(\sum_{k=0}^{\infty} (\phi_Y(t))^k \mathbb{P}(W = k) - \mathbb{P}(W = 0) \right) \\ &= p G_W(\phi_Y(t)) - p \mathbb{P}(X = 0). \end{aligned}$$

Similarly, we have

$$\begin{aligned} \mathbb{E} \left(e^{-it \sum_{i=1}^{-X} Y_i} \mathbf{1}_{\{X < 0\}} \right) &= \mathbb{E} \left(\mathbf{1}_{\{X < 0\}} \mathbb{E} \left(e^{-it \sum_{i=1}^{-X} Y_i} \mid X \right) \right) \\ &= \sum_{k=-\infty}^{-1} (\phi_Y(-t))^{-k} \mathbb{P}(X = k) \\ &= (1-p) \sum_{k=-\infty}^{-1} (\phi_Y(-t))^{-k} \mathbb{P}(W = -k) \\ &= (1-p) \left(\sum_{k=0}^{\infty} (\phi_Y(-t))^k \mathbb{P}(W = k) - \mathbb{P}(W = 0) \right) \\ &= (1-p) G_W(\phi_Y(-t)) - (1-p) \mathbb{P}(X = 0). \end{aligned}$$

Hence

$$\phi_{F \circ X}(t) = p G_W(\phi_Y(t)) + (1-p) G_W(\phi_Y(-t)). \quad (29)$$

Combining (27), (28) and (29), we obtain

$$\phi_{\varepsilon}(t) = \frac{p \phi_W(t) + (1-p) \phi_W(-t)}{p G_W(\phi_Y(t)) + (1-p) G_W(\phi_Y(-t))}, \quad t \in \mathbb{R}.$$

Finally, from the definition of the characteristic function, $\mathbb{P}(\varepsilon = k)$ represents the k th Fourier coefficient, and then (9) follows. Theorem 2 is proved. ■

Proof of Theorem 3

We have, for any $k \in \mathbb{Z}$,

$$\mathbb{P}(\varepsilon = k) = \sum_{m \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \mathbb{P}(\varepsilon_t = k, X_t = j, X_{t-1} = m) = P(k) + Q(k),$$

where

$$P(k) = \sum_{j \in \mathbb{Z}} \mathbb{P}(\varepsilon_t = k, X_t = j, X_{t-1} = 0)$$

and

$$Q(k) = \sum_{m \in \mathbb{Z}^*} \sum_{j \in \mathbb{Z}} \mathbb{P}(\varepsilon_t = k, X_t = j, X_{t-1} = m).$$

Let $\Gamma(a, b) = \mathbb{P}(X_t = a \mid X_{t-1} = b)$, $(a, b) \in \mathbb{Z}^2$, be the transition probability function of the process (3).

Observe that, thanks to (25) and the stochastic representation of the model (3), (with $X \stackrel{d}{=} X_t$)

$$P(k) = \mathbb{P}(X_t = k, X_{t-1} = 0) = \Gamma(k, 0)\mathbb{P}(X = 0) = \mathbb{P}(\varepsilon = k)\mathbb{P}(X = 0).$$

Therefore

$$\mathbb{P}(\varepsilon = k) = \frac{1}{1 - \mathbb{P}(X = 0)} Q(k).$$

Now remark that

$$\begin{aligned} Q(k) &= \sum_{m \in \mathbb{Z}^*} \sum_{j \in \mathbb{Z}} \mathbb{P} \left(X_t - \text{sign}(X_{t-1}) \sum_{i=1}^{|X_{t-1}|} Y_i = k \mid X_t = j, X_{t-1} = m \right) \times \\ &\quad \mathbb{P}(X_t = j, X_{t-1} = m) \\ &= \sum_{m \in \mathbb{Z}^*} \sum_{j \in \mathbb{Z}} \mathbb{P} \left(\text{sign}(m) \sum_{i=1}^{|m|} Y_i = j - k \right) \Gamma(j, m) \mathbb{P}(X = m) \\ &= Q_1(k) + Q_2(k), \end{aligned}$$

where, by (5) and (26), for $Z_m \sim \text{Binomial}(2m, \alpha)$,

$$\begin{aligned} Q_1(k) &= \sum_{m=1}^{\infty} \mathbb{P}(X = m) \sum_{j \in \mathbb{Z}} \mathbb{P} \left(\sum_{i=1}^m Y_i = j - k \right) \Gamma(j, m) \\ &= \sum_{m=1}^{\infty} \mathbb{P}(X = m) \sum_{j=k-m}^{k+m} \mathbb{P}(Z_m = m + j - k) \Gamma(j, m) \\ &= \sum_{m=1}^{\infty} \mathbb{P}(X = m) \sum_{j=k-m}^{k+m} \sum_{l=-m}^m \mathbb{P}(\varepsilon = j - l) \mathbb{P}(Z_m = m + j - k) \mathbb{P}(Z_m = m + l) \end{aligned}$$

and

$$\begin{aligned}
Q_2(k) &= \sum_{m=-\infty}^{-1} \mathbb{P}(X = m) \sum_{j \in \mathbb{Z}} \mathbb{P} \left(\sum_{i=1}^{-m} Y_i = -(j - k) \right) \Gamma(j, m) \\
&= \sum_{m=1}^{\infty} \mathbb{P}(X = -m) \sum_{j \in \mathbb{Z}} \mathbb{P} \left(\sum_{i=1}^m Y_i = -(j - k) \right) \Gamma(j, -m) \\
&= \sum_{m=1}^{\infty} \mathbb{P}(X = -m) \sum_{j=k-m}^{k+m} \mathbb{P}(Z_m = m - j + k) \Gamma(j, -m) \\
&= \sum_{m \in \mathbb{N}^*} \mathbb{P}(X = -m) \sum_{j=k-m}^{k+m} \sum_{l=-m}^m \mathbb{P}(\varepsilon = j - l) \mathbb{P}(Z_m = m - j + k) \mathbb{P}(Z_m = m - l).
\end{aligned}$$

Theorem 3 is proved. ■

Proof of Theorem 4

Implicit form for the stationary distribution. By the stochastic representation of the model (3) and the independence between ε_t and the sigma-algebra $\mathcal{G}_{t-1} = \sigma(X_{t-1}, \dots)$, for any $k \in \mathbb{Z}$, we have (with $X \stackrel{d}{=} X_t$)

$$\begin{aligned}
\mathbb{P}(X = k) &= \sum_{j \in \mathbb{Z}} \mathbb{P}(F_1 \circ X = j) \mathbb{P}(\varepsilon = k - j) \\
&= \mathbb{P}(F_1 \circ X = 0) \mathbb{P}(\varepsilon_1 = k) + \sum_{j \in \mathbb{Z}^*} \mathbb{P}(F_1 \circ X = j) \mathbb{P}(\varepsilon = k - j).
\end{aligned}$$

Calculation of $F_1 \circ X$ distribution. For all $l \in \mathbb{Z}$, we have

$$\begin{aligned}
\mathbb{P}(F_1 \circ X = l) &= \sum_{k \in \mathbb{Z}} \mathbb{P}(X = k, F_1 \circ X = l) \\
&= \sum_{k \in \mathbb{Z}} \mathbb{P}(F_1 \circ X = l \mid X = k) \mathbb{P}(X = k) \\
&= \sum_{k \in \mathbb{Z}} \mathbb{P} \left(\text{sign}(k) \sum_{i=1}^{|k|} Y_i = l \right) \mathbb{P}(X = k) \\
&= A(l) + B(l) + C(l),
\end{aligned}$$

where

$$A(l) = \mathbb{P}(X = 0) \mathbf{1}_{\{l=0\}},$$

by (5), for $Z_k \sim \text{Binomial}(2k, \alpha)$,

$$\begin{aligned} B(l) &= \sum_{k=1}^{\infty} \mathbb{P} \left(\sum_{i=1}^k Y_i = l \right) \mathbb{P}(X = k) \\ &= \sum_{k=1}^{\infty} \mathbb{P}(Z_k = k + l) \mathbb{P}(X = k) \mathbf{1}_{\{l \in \{-k, \dots, k\}\}} \end{aligned}$$

and

$$\begin{aligned} C(l) &= \sum_{k=-\infty}^{-1} \mathbb{P} \left(\sum_{i=1}^{-k} Y_i = -l \right) \mathbb{P}(X = k) \\ &= \sum_{k=1}^{\infty} \mathbb{P} \left(\sum_{i=1}^k Y_i = -l \right) \mathbb{P}(X = -k) \\ &= \sum_{k=1}^{\infty} \mathbb{P}(Z_k = k - l) \mathbb{P}(X = -k) \mathbf{1}_{\{l \in \{-k, \dots, k\}\}}. \end{aligned}$$

Combining the values of $A(l)$, $B(l)$ and $C(l)$, we end the proof of Theorem 4. ■

Proof of Theorem 5

Note that we consider a P-SINAR(1) process, as defined in (3). Thus, the existence of a unique stationary solution, denoted by X , and the ergodicity of the process are ensured. Therefore, consistency and asymptotic normality of conditional least squares estimator follow from the result of Klimko and Nelson (1978), as it is explained in Du and Li (1991) for the case of INAR processes. Now, it remains to calculate the finite variance Ψ of the normal distribution in (16), under the assumption $X \sim \text{SPo}(\lambda)$, i.e. $X \stackrel{d}{=} RW$, where $R \sim \text{Rademacher}(1/2)$ and $W \sim \text{Poisson}(\lambda)$. Hence, we note that

$$\sqrt{n}(\hat{\alpha}_n - \alpha_1) = \frac{n^{-\frac{1}{2}} \sum_{t=1}^n X_{t-1}(X_t - \alpha_1 X_{t-1})}{n^{-1} \sum_{t=1}^n X_{t-1}^2}$$

and further that the denominator converges almost surely to $\mathbb{E}(X^2) = \lambda(1+\lambda)$. So, by the law of large numbers and the continuous mapping theorem, this has the same asymptotic distribution as

$$\frac{1}{\lambda(1+\lambda)} n^{-\frac{1}{2}} \sum_{t=1}^n X_{t-1}(X_t - \alpha_1 X_{t-1}).$$

Since this sum is a martingale and martingale differences are stationary ergodic with finite moments, it converges to a normal distribution with mean zero and

variance

$$\begin{aligned}
\Psi &= \frac{1}{\lambda^2(1+\lambda)^2} \mathbb{E}(X_{t-1}^2 (X_t - \alpha_1 X_{t-1})^2) \\
&= \frac{1}{\lambda^2(1+\lambda)^2} \mathbb{E}(X_{t-1}^2 \mathbb{V}(X_t | X_{t-1})) \\
&= \frac{1}{\lambda^2(1+\lambda)^2} (\beta_1 \mathbb{E}(X_{t-1}^2 | X_{t-1}) + \sigma_\varepsilon^2 \mathbb{E} X_{t-1}^2) \\
&= \frac{1}{\lambda^2(1+\lambda)^2} (\beta_1 \mathbb{E}(X^2 | X) + \sigma_\varepsilon^2 \mathbb{E}(X^2)),
\end{aligned}$$

where $\alpha_1 = 2\alpha - 1$ and $\beta_1 = 2\alpha(1 - \alpha)$ denotes respectively the mean and variance of F_1 , and σ_ε^2 the variance of ε . Now, from (8), we have

$$\begin{aligned}
\sigma_\varepsilon^2 &= (1 - \alpha_1^2) \mathbb{V}(X) - \beta_1 \mathbb{E}(|X|) \\
&= \beta_1 (2\mathbb{V}(X) - \mathbb{E}(|X|)).
\end{aligned}$$

Recall that $W \sim \text{Poisson}(\lambda)$, and

$$\mathbb{V}(X) = \mathbb{E}(X^2) = \mathbb{E}(W^2) = \lambda(1 + \lambda).$$

Thus

$$\mathbb{E}(|X|) = \mathbb{E}(W) = \lambda,$$

and then

$$\sigma_\varepsilon^2 = 2\alpha\lambda(1 - \alpha)(2\lambda + 1).$$

Note also that

$$\mathbb{E}(X^2 | X) = \mathbb{E}(W^3) = \lambda(\lambda^2 + 3\lambda + 1).$$

Therefore,

$$\begin{aligned}
\Psi &= \frac{1}{\lambda^2(1+\lambda)^2} (\beta_1 \mathbb{E}(W^3) + \sigma_\varepsilon^2 \mathbb{E}(W^2)) \\
&= 2\alpha(1 - \alpha) \frac{2\lambda^3 + 4\lambda^2 + 4\lambda + 1}{\lambda(1 + \lambda)^2}.
\end{aligned}$$

Theorem 5 is proved. ■

Annex: on the behavior of a R-Po(p, λ) distribution with large λ

In the sequel, by sake of simplicity, we suppose that $\lambda = n \in \mathbb{N}^*$. Let $p \in (0, 1)$ and $X_n = RW_n$, where R and W_n are two independent random variables such

that $R \sim \text{Rademacher}(p)$ and $W_n \sim \text{Poisson}(n)$. Hence, for any $x \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{P}(X_n \leq x) &= (1-p)\mathbb{P}(-W_n \leq x) + p\mathbb{P}(W_n \leq x) \\ &= (1-p)(1 - \mathbb{P}(W_n \leq -x)\mathbf{1}_{\{x \leq 0\}}) + p\mathbb{P}(W_n \leq x)\mathbf{1}_{\{x \geq 0\}} \\ &= (1-p) \left(1 - \mathbb{P}\left(\frac{W_n}{\sqrt{n}} - \sqrt{n} \leq -\frac{x}{\sqrt{n}} - \sqrt{n}\right) \mathbf{1}_{\{x \leq 0\}} \right) \\ &\quad + p\mathbb{P}\left(\frac{W_n}{\sqrt{n}} - \sqrt{n} \leq \frac{x}{\sqrt{n}} - \sqrt{n}\right) \mathbf{1}_{\{x \geq 0\}}. \end{aligned}$$

By the central-limit theorem, for n large enough, one can approximate $W_n/\sqrt{n} - \sqrt{n}$ by a random variable $U \sim \mathcal{N}(0, 1)$. Now, let f_U be the density function of U and F_U its cumulative distribution function. Therefore,

$$\begin{aligned} \mathbb{P}(X_n \leq x) &\approx (1-p) \left(1 - F_U\left(-\frac{x}{\sqrt{n}} - \sqrt{n}\right) \mathbf{1}_{\{x \leq 0\}} \right) \\ &\quad + pF_U\left(\frac{x}{\sqrt{n}} - \sqrt{n}\right) \mathbf{1}_{\{x \geq 0\}} \\ &= G_n(x), \end{aligned}$$

where $G_n(x)$ denotes a cumulative distribution function of density

$$\begin{aligned} g_n(x) = G_n(x)' &= (1-p) \frac{1}{\sqrt{n}} f_U\left(-\frac{x}{\sqrt{n}} - \sqrt{n}\right) \mathbf{1}_{\{x \leq 0\}} \\ &\quad + p \frac{1}{\sqrt{n}} f_U\left(\frac{x}{\sqrt{n}} - \sqrt{n}\right) \mathbf{1}_{\{x \geq 0\}}, \quad x \in \mathbb{R}. \end{aligned}$$

Thus, one can deduce that with large values for λ , a $\text{R-Po}(p, \lambda)$ distribution can be sufficiently approximated by two gaussian distributions: one on \mathbb{N} and the other on $\mathbb{Z}^- = \mathbb{Z} \setminus \mathbb{N}$.

In Figure 7, we depict the probability mass function of a $\text{R-Po}(1/2, \lambda)$ distribution (called also Symmetric-Poisson distribution and denoted by $\text{SPo}(\lambda)$) for $\lambda \in \{4, 6, 8, 10\}$. For each case, the blue curve represents the function $g_n \equiv g_\lambda$ given above.

Figures

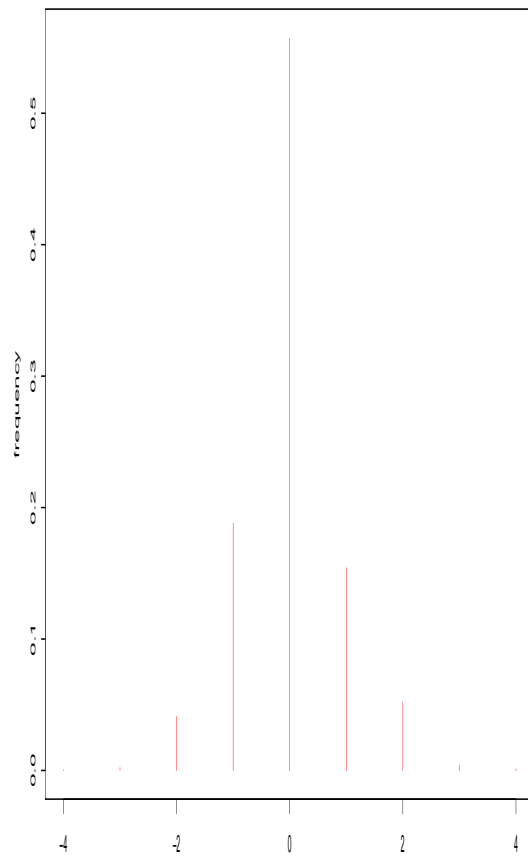


Fig. 1 The probability distribution of the noise ε , when X follows the distribution defined by (24).

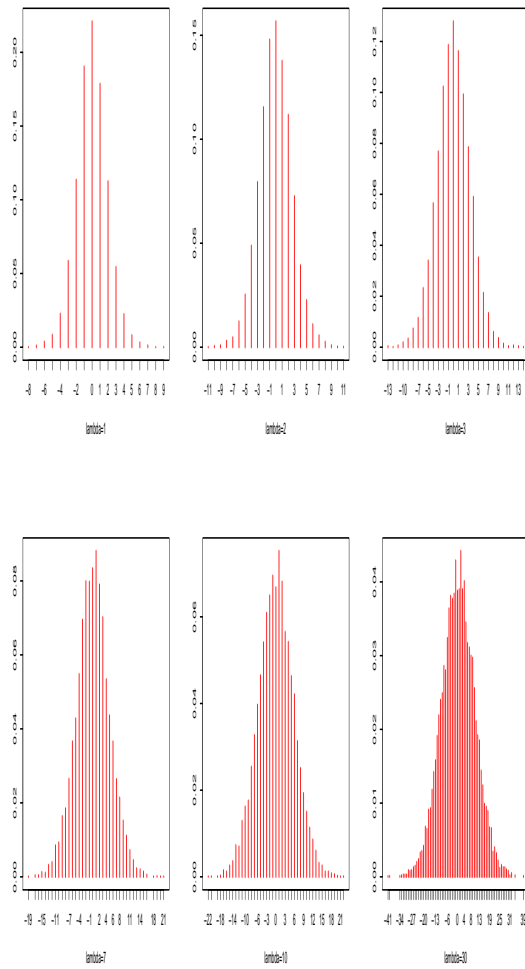


Fig. 2 The probability stationary distribution X , when $\varepsilon \sim \text{Skellam}(\lambda)$ with different values of λ .

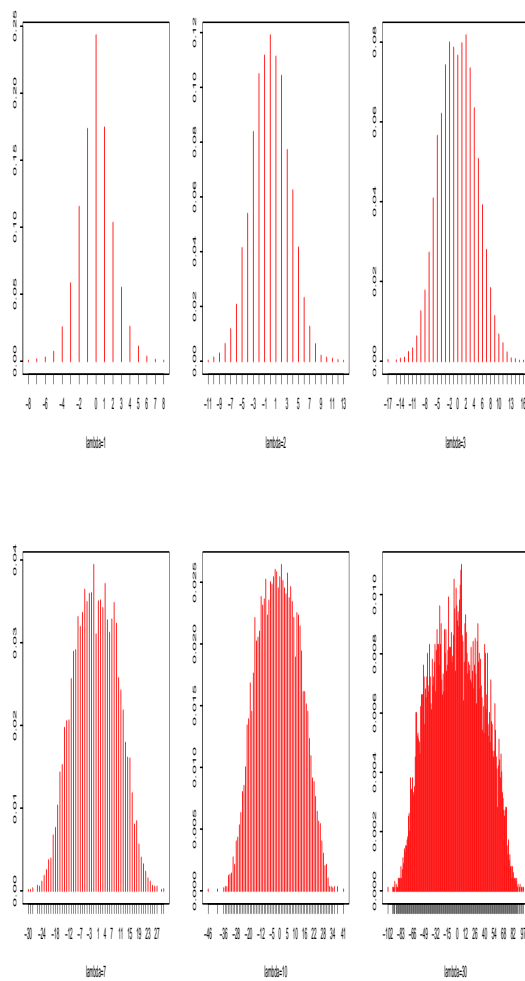


Fig. 3 The probability stationary distribution X , when $\varepsilon \sim \text{SPo}(\lambda)$ with different values of λ .

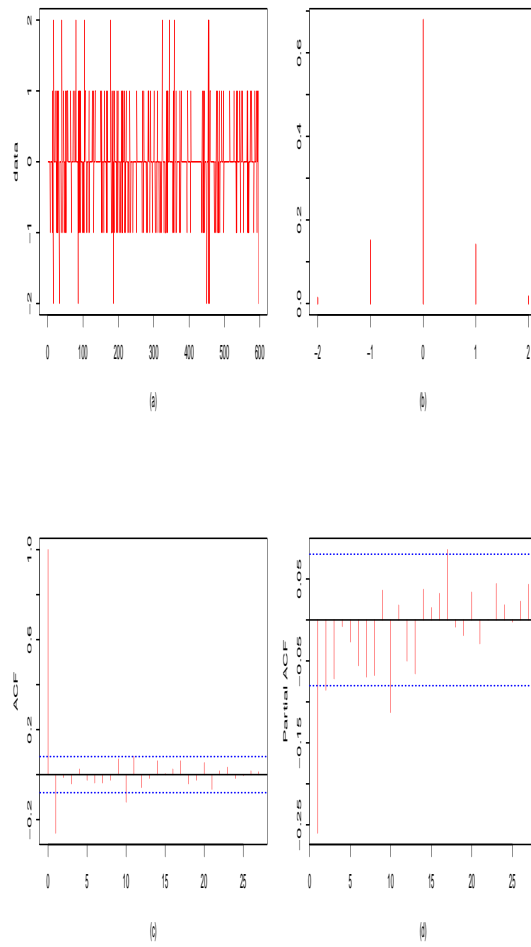


Fig. 4 (a) Time series plot of the data. (b) Histogram of the data. (c) Sample autocorrelation function. (d) Sample partial autocorrelation function.

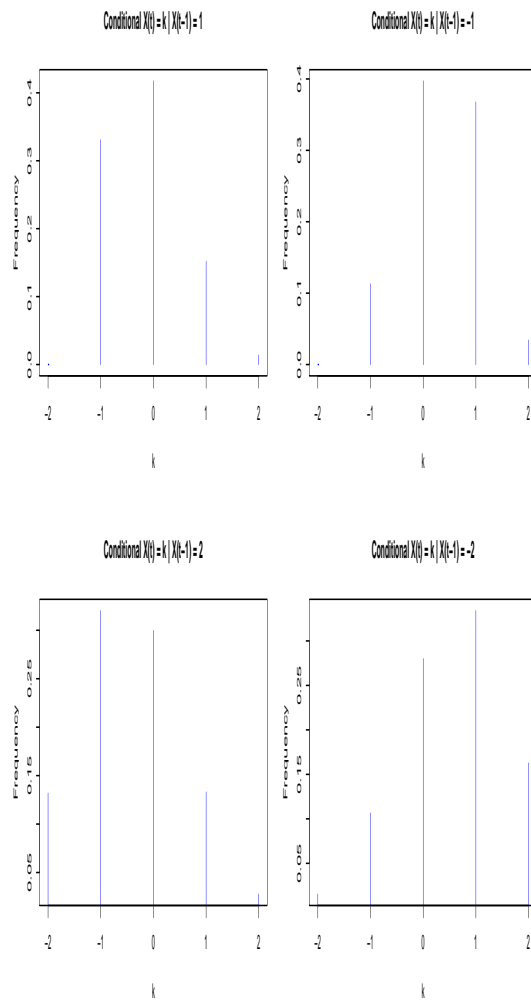


Fig. 5 Predictive distributions for $X_t \mid X_{t-1} = i$ for values of $i = 1, -1, 2, -2$, with $\varepsilon \sim \text{SPo}(\lambda)$.

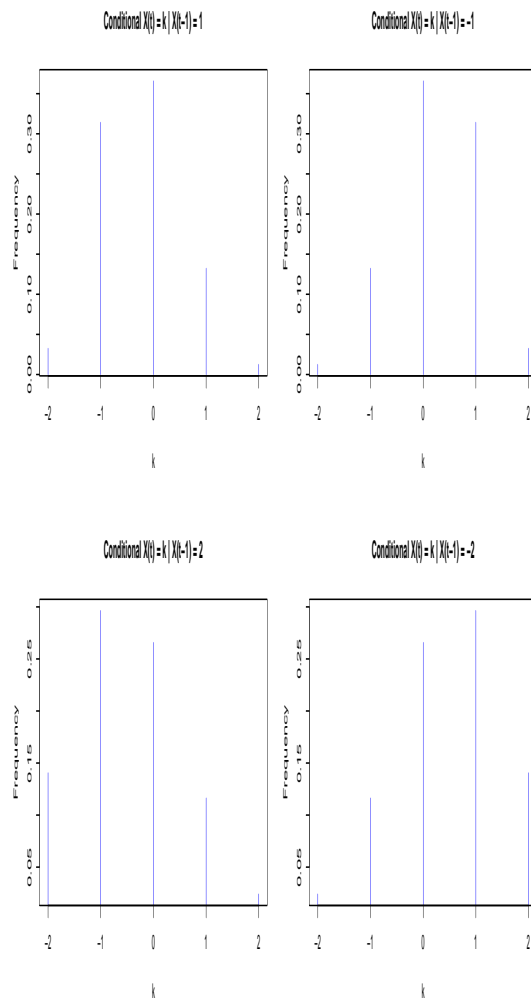


Fig. 6 Predictive distributions for $X_t \mid X_{t-1} = i$ for values of $i = 1, -1, 2, -2$, with $\varepsilon \sim \text{Skellam}(\lambda)$.

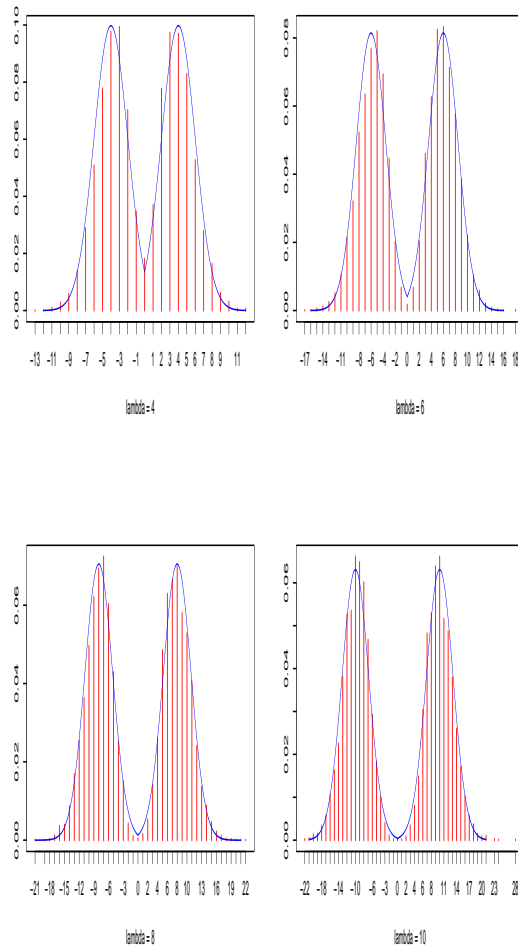


Fig. 7 The probability distribution of a $\text{SPo}(\lambda)$ with different values of λ .

References

- Al-zaid, A.A. and Al-Osh, M. (1987). First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.*, 8, (3), 261-275.
- Al-zaid, A.A. and Al-Osh, M. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *J. Appl. Probab.*, 27, (2), 314-324.
- Al-Zaid, A.A. and Omair, M.A. (2010). On the poisson difference distribution inference and applications. *Bulletin of the Malaysian Mathematical Sciences Society*, 8, (33), 17-45.
- Du, J.-G. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *J. Times Ser. Anal.*, 12, 129-142.
- Freeland, R.K. (2010). True integer value time series. *Advances in Statistical Analysis*, 94, 217-229.
- Gauthier, G. and Latour, A. (1994). Convergence forte des estimateurs des paramètres d'un processus GENAR(p). *Ann. Sci. Math. Qubec*, 18, (1), 49-71.
- Inusah S. and Kozubowski, T.J. (2006). A skew Laplace distribution on integers. *Annals of the Institute of Statistical Mathematics*, 58, (3), 555-571.
- Jung, R.C. and Tremayne, A.R. (2006). Binomial thinning models for integer time series. *Stat. Model.*, 6, (2), 81-96.
- Kachour, M. (2011). On the rounded integer-valued autoregressive process. Preprint.
- Kachour, M. and Truquet, L. (2011). A p-order signed integer-valued autoregressive (SINAR(1)) model. *Journal of Time Series Analysis*, 32, (4), 223-236.
- Kachour, M. and Yao, J.F. (2009). First-order rounded integer-valued autoregressive (RI-NAR(1)) process. *Journal of Time Series Analysis*, 30, (4), 417-448.
- Karlis, D. and Ntzoufras, I. (2006). Bayesian analysis of the differences of count data. *Statistics in medicine*, 25, (11), 1885-1905.
- Karlis, D. and Ntzoufras, I. (2009). Bayesian modeling of football outcomes: Using the skellam distribution for the goal difference. *Journal Management Mathematics*, 20, (2), 133-145.
- Katti, S.K. (1960). Moments of the Absolute Difference and Absolute Deviation of Discrete Distributions. *Annals of Statistics*, 31, 78-85.
- Kedem, B. and Fokianos, K. (2002). Regression models for time series analysis. Wiley-Interscience New-York.
- Kemp, A.W. (1997). Characterizations of a discrete normal distribution. *Journal of Statistical Planning and Inference*, 63, (2), 223-229.
- Kim, H.-Y. and Park, Y. (2008). A non-stationary integer-valued autoregressive model. *Statistical papers*, 49, (3), 485-502.
- Klimko, L.A. and Nelson, P.I. (1978). On conditional least squares estimation for stochastic process. *Ann. Statist.*, 6, 629-642.
- Latour, A. and Truquet, L. (2008). An integer-valued bilinear type model. Preprint.
- McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin*, 21, (4), 645-650.
- McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Adv. in Appl. Probab.*, 20, 822-835.
- McKenzie, E. (2003). Discrete variate time series. In Stochastic processes: modeling and simulation, volume 21 of Handbook of Statist., pages 573-606. North- Holland, Amsterdam.
- Ong, S.H., Shimizu, K. and Min Ng, C. (2008). A class of discrete distributions arising from difference of two random variables. *Computational Statistics and Data Analysis*, 52, (3), 1490-1499.
- Ristic, M.M., Bakouch, H.B. and Nastic, A.S. (2009). A new geometric first-order integer-valued autoregressive(NGINAR(1))process. *Journal of Statistical Planning and Inference*, 139, 2218-2226.
- Ristic, M.M., Bakouch, H.B. and Nastic, A.S. (2009). Zero truncated poisson integer-valued AR(1) model. *Metrika*, 72, (2), 265-280.
- Roy, D. (2003). The discrete normal distribution. *Communications in Statistics - Theory and Methods*, 32, (10), 1871-1883.

- Skellam, J.G. (1946). The frequency distribution of the difference between two poisson variates belonging to different populations. *Journal of the Royal Statistical Society, Series A*, 109-296.
- Steutel, F.W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Ann. Probab.*, 7, (5), 893-899.
- Szablowski, P.J. (2001). Discrete normal distribution and its relationship with jacobi theta functions. *Statistics and Probability Letters*, 52, (3), 289-299.
- Wei, C.H. (2008). Thinning operations for modeling time series of counts a survey. *J. Appl. Probab.*, 92, (3), 319-341.
- Zhang, H., Wang, D. and Zhu, F. (2010). Inference for INAR(p) processes with signed generalized power series thinning operator. *Journal of Statistical Planning and Inference*, 140, (3), 667-683.